DOCUMENT RESUME

ED 292 814 TM 011 183

AUTHOR Barrios, Nina B.; Foster, Garrett R.
TITLE Utilization of Evaluation Information: A Case Study Approach Investigating Factors Related to Evaluation Utilization in a Large State Agency.
PUB DATE Oct 87
NOTE 34p.; Paper presented at the Annual Meeting of the American Evaluation Association (Boston, MA, October 14-17, 1987).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Research/Technical (143) -- Tests/Evaluation Instruments (160)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Case Studies; Evaluation Methods; *Evaluation Utilization; *Information Utilization; Program Evaluation; *Scaling; State Agencies; Statistical Analysis; Use Studies
IDENTIFIERS Equal Interval Scoring

ABSTRACT

The use of evaluation information in a large state agency was investigated through a case study approach. Utilization was measured through a scaling procedure considering the status of recommendations within a report and the extent of influence of the evaluation of the status. The scaling procedure, the equal-appearing interval method, used judgments of 15 experienced evaluators to obtain scale values for the relative importance of varying combinations of status and influence in the overall assessment of the effectiveness of evaluation information. The scale was tested on 11 reports from a state human services agency. Four of these were selected for in-depth case study. The contrast of one high utilization and three low utilization situations gave a basis for assessing the value of predictor factors. The scale appeared to function as intended, discriminating between varying levels of utilization. Findings also indicate that the content of evaluation information is an important factor in utilization. Five appendices include statistical data, the survey instrument, category definitions and response form, and the organization of variables for interviews and analysis. (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED292814

UTILIZATION OF EVALUATION INFORMATION:

A case study approach investigating factors related to evaluation utilization in a large state agency

by

Nina B. Barrios, Ph.D.
Evaluation Systems Design, Inc.

and

Garrett R. Foster, Ph.D.
Professor, Florida State University

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

✓ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

NINA B. BARRIOS

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Presented at the October 1987 meeting of
American Evaluation Association

TM 011 183

BEST COPY AVAILABLE

Utilization of evaluation information: A case study approach investigating factors related to evaluation utilization in a large state agency

Introduction

Early studies of evaluation utilization were characterized by imprecise definition of concepts as well as inadequate and undocumented measurement strategies. This state of affairs resulted in contradictory findings, debates over conclusions and varying recommendations regarding the methods that would lead to productive study of the factors that explain utilization of evaluation information. In the steadily increasing literature on the utilization of social research and evaluation, there now appears to be a consensus forming on (1) the issue of whether information produced by evaluations is used, (2) how the concept of utilization is defined, and (3) a preliminary identification of various factors that affect the extent to which evaluation information is used. Even though there is diminished debate over whether use occurs and what major factors affect it, there remain significant theoretical and methodological shortcomings to the study of utilization.

This research attempted to address these shortcomings through the study of utilization in its natural context and through the use of a unique measurement strategy. The study used an in-depth case study approach to look at the usefulness of variables that many have identified as relevant to utilization, including the content of recommendations in evaluation reports. The case studies contrasted evaluation reports where the extent of utilization was high and where utilization was low. To identify the high-use and low-use cases, a measurement strategy was employed that yielded an assessment of utilization that is comparable across studies, and one which incorporated a multidimensional measure of use. There were, therefore, two primary objectives in this research. The first objective concerned methodological problems, specifically the **measurement of utilization**. The second objective was to investigate factors relating to utilization through **case studies** of high-use and low-use reports.

Measurement of utilization

Inadequate methods for measuring utilization have limited the development of utilization theory. The measurement of use in most studies reflects the simplest and least informative methods such as a simple count of instances of "use" or identifying the percentage of recommendations implemented. In particular, past measurement strategies have seldom taken into account such things as the competing influence of other sources of information, cost of implementation of recommendations, organizational and second party requirements for implementation, etc. This study is based on the proposition that nominal measures and percentages are not adequate to represent the complex nature of utilization of evaluation information. In addition to identifying specific inadequacies of conventional methods for measuring utilization, this study presents a measurement strategy that addresses many of the current shortcomings. The measurement strategy developed in this research yields scale values for evaluation reports, enabling a comparative rating of reports in terms of utilization. Qualitative information is used to augment and validate the results of the scaling procedure.

Case studies

The case study approach was used to investigate in context those factors that relate to utilization of evaluation in a large state agency. The case studies investigated four evaluation studies where the extent of utilization was high and three where utilization was low. The contrast of high-use and low-use reports provided a basis for assessing the potency of various utilization predictors (factors) reported in the literature. The research focused on two questions: (1) To what extent was evaluation information utilized in a large state social service agency? and (2) What conditions best differentiated between high and low levels of utilization of evaluation information? Concerning the latter question, the study investigated the contribution of a variety of technical and organizational variables supported in the literature as enhancing utilization. The study also considered an area not generally investigated, that is, the content of evaluation recommendations. Through the case studies, the study resulted in identification of a composite of characteristics of evaluation reports and the evaluation process that were related to utilization of evaluation information.

Status of measurement of utilization

Initially, the development of measurement methods was overshadowed by the debate over what to measure, i.e., what constitutes use. Early studies of utilization were based on a restricted, behavioral definition of utilization where "immediate and direct" use was the criterion (Weiss, 1972). Studies in the late 70's and early 80's expanded the concept of utilization to include cognitive as well as behavioral uses (Patton, 1977; Weiss and Bucuvalas, 1980; and Rich, 1981). Presently, the initial, limiting definition of utilization has been unanimously discarded. Researchers now acknowledge a broader definition incorporating three parts: instrumental use (action or planned action based on evaluation information), conceptual use (influence of decision-maker's thinking about a program or issue), and persuasive use (where evaluation information is used by decision-makers to influence others to a particular point of view) (Weiss 1981; Braskamp and Brown, 1980; Heilman, 1983). Most of the current literature on utilization reflects a broader definition of utilization that includes conceptual and persuasive uses as legitimate and desirable applications of evaluation information.

A consensus on the definition of utilization has not produced uniform measurement of the concept. Virtually every study uses different measurement strategies. Most measurement strategies in studies of utilization are variations of: (1) rating scales and (2) interviews of decision-makers. Rating scales largely look at an evaluation report as a whole and identify a nominal category or one point on an ordinal scale to represent the extent to which evaluation information was used. Interviews are typically used as part of a case study approach, which currently is the most frequent research method for utilization studies. Interview strategies generally identify the primary audiences of an evaluation study and, from these individuals, identify additional users or secondary audiences.

Both of the general methods used in utilization studies have positive and negative aspects. Both rating scales and interviews have numerous attributes, positive and negative, that are idiosyncratic to the developer and the reader. Perhaps the most serious limitation of current utilization assessment is the fact that they do not take into account the competition of other influences on decision-makers' actions. For example, an action can be taken, <u>not</u> because it was recommended in an evaluation, but because it was

heralded by an influential political figure quite independently of the evaluation. In this case, although the recommended action was taken, the evaluation may have had little influence on the decision-maker. Rating scales, while easily quantified, seldom take into account the conceptual and persuasive uses of evaluation information. Interviews, on the other hand, are well suited to identifying competing influences and exploring conceptual and persuasive use. Interview data, of course, are much more difficult to quantify and compare. Further, published studies using rating scales and interviews rarely present sufficient documentation of the forms or questions used in such a way that the study could be replicated by other researchers.

A critical review of current literature indicates that a desirable method for investigating utilization would: (1) incorporate the three-part definition of utilization presently accepted in the literature (instrumental, conceptual, and persuasive), (2) look at instances of recommended and unprompted actions, (3) account for competing influence on decision-makers, and (4) providing a set of operation that facilitates replication and is widely applicable.

Development of measurement strategy

The focus of this study was utilization of the information contained in eleven evaluation reports prepared by the internal evaluation unit of a large state agency. The measurement of utilization incorporated information from primary and secondary audiences of the evaluation reports. From each group, information was obtained about instrumental use as well as conceptual, persuasive, and other uses. The following discussion describes the development and implementation of the strategy tested in this research to provide quantitative and qualitative information with which to measure utilization. Table 1, presented at the end of the discussion, provides a summary of the development and implementation process.

Scale development

The method for assessing instrumental utilization was a scaling procedure which integrated the concepts of status of evaluation recommendation implementation and the extent of influence of the evaluation on status. The categories representing the status of evaluation recommendations are listed below and defined in Appendix A:

1. implemented as stated
2. implemented with modifications
3. planned implementation as stated
4. planned implementation with modifications
5. not to be implemented
6. decision not yet made

The categories representing the extent of influence of the evaluation information on the decision were:

1. predominant influence
2. major influence
3. moderate influence
4. minor influence
5. no influence

The scaling procedure selected was the equal-appearing interval method as described by Guilford (1954). This method uses comparative judgments to obtain scale values for statements. For this study, all logical pairs of status and influence were listed. Raters were solicited to review the list of paired items and indicate, on a scale of one to 11, their judgment of the value relative to utilization. The median value of the distribution of judgments for each item was taken as the scale value. The interquartile range or Q-value was calculated to measure the dispersion or variation of responses to an item. The scale was administered to 15 respondents, all experienced evaluators of state programs. The scale was administered in a group setting with no discussion of individual judgments prior to the ratings. Instructions included specific clarification of several items where the pilot had indicated a potential for confusion. The scale and instructions are included in Appendix A.

Overall, the scale appeared to function as intended, providing meaningful scale values that discriminate between the varying levels of status and influence. Most Q values indicated an acceptable amount of dispersion. In a few instances, there was more dispersion than was desired, but the situations were not judged to be sufficiently weak to merit exclusion from the scale. The scale values and Q values for each item are contained in Appendix B.

Application of scale

For each of the 11 reports, the primary audience was asked to indicate the status of each recommendation according to the categories of implementation and influence. A form was developed for each of the 11 reports which contained the recommendations and a space to record judgments. An example of the response form is contained in Appendix C. In each case, the individual representing the primary audience of the report agreed to complete the form.

Several methods of determining the reliability of the ratings were considered. Ideally, multiple ratings would have been obtained to allow judgment of inter-rater reliability; however, circumstances were such that each rater was in a unique position within the organization in terms of level of authority and program context. Any attempt to duplicate the rating would have involved an individual, either a superior or subordinate, with different responsibilities and therefore different knowledge and perceptions of the issue. In each case, the rater was the primary decision-maker and the one most able to make an accurate, informed rating. Therefore, an intra-rater procedure was used to determine reliability. To obtain the initial ratings, a face-to-face meeting was held with each respondent where the rating categories were explained and questions addressed. Confidentiality was assured and assurance was given that the ratings would not be interpreted as a judgment of the quality of management. None of the raters expressed or displayed any concern over making the ratings. About one month after the ratings were obtained, the same individuals who initially rated the seven reports selected for the study were asked to review their ratings and determine if any changes were appropriate after the "second reading". Only one of the seven individuals made any changes that resulted in an adjustment to the rank order of the report scores. With this exception, the rank order and groupings of the studies remained the same.

Qualitative information from primary audience

In addition to the ratings reflecting instrumental uses of evaluation information, members of the primary audience were also asked to give information about:

(1) pre-emptive use of evaluation information (decisions made before the evaluation recommendations were published, but that were influenced by the evaluation study,
(2) conceptual uses of evaluation (information that was enlightening or educational),
(3) persuasive use (use of information to persuade others to a particular point of view),
(4) incidental use (action taken as a result of evaluation information though not specifically recommended in the report, and
(5) other uses as identified by the recipient.

A questionnaire was developed which reflected the forms of utilization listed above. (See Appendix D.)

Data collection from secondary audiences

A set of questions dealing with direct and indirect uses of evaluation information was compiled into a survey and mailed to secondary audiences. One question focused on instrumental uses, while the others dealt with conceptual and persuasive forms of utilization. Other than the question on instrumental use, the questions were identical to those presented to primary audiences. The survey was sent to 197 individuals or agencies, identified by the evaluators who conducted the study, as recipients of one of the 11 evaluation reports.

Table 1 presents a summary of the development and implementation process investigated by this research for the measurement of utilization.

Results of the measurement process

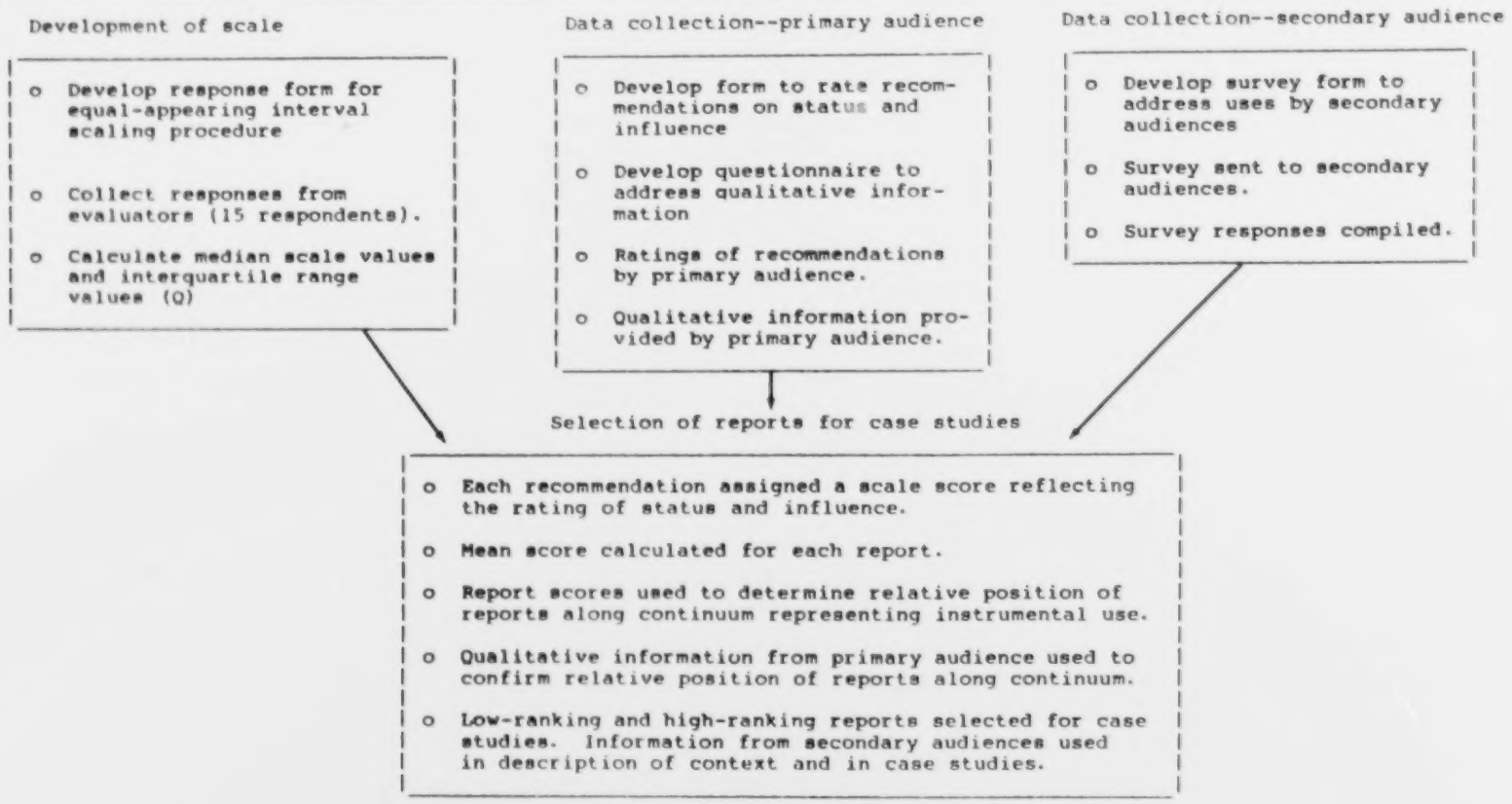
Use of evaluation information by the agency

Based on the sample of 11 reports, a general description of the use of evaluation information by the agency was developed. The description addresses instrumental use by the primary and secondary audiences, and conceptual and persuasive use by both groups.

Instrumental use. A total of 126 recommendations was contained in the 11 reports used in this study, ranging from a low of three recommendations in one study to a high of 20 recommendations in another. The mean number of recommendations per report was 11.5; the median was 12 and the mode was 15. Examining the separate categories of implementation status and influence, there appears to be a great deal of instrumental use of evaluation information, as indicated by the high percentage of recommendations implemented and, more importantly, by the percentage of recommendations where the evaluation is credited as being a predominant or major influence, or at least a moderate influence. Eighty percent of the recommendations were implemented or to be implemented, either as written or with modifications. Only 20 percent were not to be implemented or had not been decided upon within six months after the publication of the report. In terms of influence, for 44 percent of the recommendations, the evaluation was credited with predominant, major, or moderate influence on the decisions. The recommendations were judged to have had minor influence in 21 percent of the recommendations. Only 25 percent of the recommendations were judged to have been of no influence. (In ten percent of the cases, no decision had been made, so influence could not be judged.)

Table 1

Summary of Development and Implementation of Measurement Strategy


Development of scale
o Develop response form for equal-appearing interval scaling procedure
o Collect responses from evaluators (15 respondents).
o Calculate median scale values and interquartile range values (Q)
Data collection--primary audience
o Develop form to rate recommendations on status and influence
o Develop questionnaire to address qualitative information
o Ratings of recommendations by primary audience.
o Qualitative information provided by primary audience.
Data collection--secondary audience
o Develop survey form to address uses by secondary audiences
o Survey sent to secondary audiences.
o Survey responses compiled.
Selection of reports for case studies
o Each recommendation assigned a scale score reflecting the rating of status and influence.
o Mean score calculated for each report.
o Report scores used to determine relative position of reports along continuum representing instrumental use.
o Qualitative information from primary audience used to confirm relative position of reports along continuum.
o Low-ranking and high-ranking reports selected for case studies. Information from secondary audiences used in description of context and in case studies.

In addition to instrumental use of evaluation information by the primary audience, this study also looked at use by secondary audiences. The survey information obtained from secondary audiences indicated that, out of 66 individuals familiar with one of the reports, 39 (59 %) indicated that they had made some instrumental use of the information.

Conceptual and persuasive use. The information in the preceding section related to the instrumental use of evaluation information. To gather information about conceptual and persuasive uses and other uses, the primary and secondary audiences were asked several specific questions in addition to the questions about instrumental use. In terms of cognitive use by primary audiences, seven of 11 (64%) reported use of the evaluation for enlightenment, i.e., enhancing their understanding of the program. Five of the 11 reported persuasive uses of the information and three of the 11 primary users reported pre-emptive use.

For secondary audiences, 82 percent of 66 respondents reported cognitive use in terms of gaining a better understanding of the program. Persuasive use was indicated by 20 percent of the secondary audience respondents.

Using scale scores to assess evaluation reports

Based on the ratings of status and influence provided by the primary audience, each recommendation in each of the 11 reports was assigned a scale value according to the results of the equal-appearing interval scaling procedure. To score reports, the scores for each recommendation were totaled and the mean and standard deviation computed for each report.

Figure 1 shows the mean score and standard deviation of each of the 11 reports (identified as A through K) and the overall mean (X). As Figure 1 illustrates, there were clearly three low-ranking and four high-ranking reports. Studies A, B, and C were identified as a "low-use" group and Studies H, I, J, and K were identified as the "high-use" group. The additional qualitative information from the primary audiences was reviewed and determined to support the classification of the seven reports in to the high- and low-use groups.

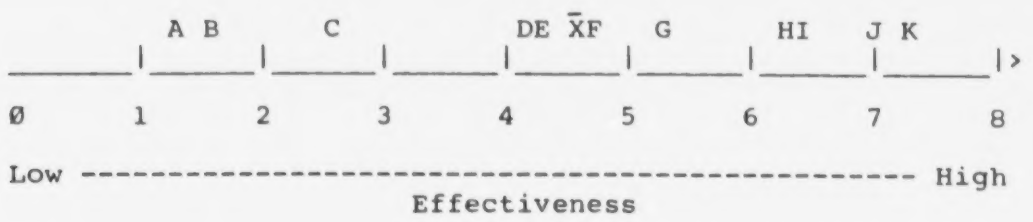
The validity claim or trustworthiness of the utilization assessment is the process of selecting the individual who provided the ratings, a "member check" of the scale values for the reports (interviewing the director of the evaluation unit conducting the studies), and the extension of the concept of instrumental use to include "influence" as well as the inclusion of cognitive and persuasive dimensions of information utilization.

The utilization literature

Five published works ranging from 1979 to 1982 have attempted to synthesize the numerous and varied studies over the past 15 years (Young and Comptois, 1979; Braskamp and Brown, 1980; Rothman, 1980; Leviton and Hughes, 1981; Beyer and Trice, 1982). A review of these studies reveals commonalities among major factors proposed by the reviewers. Appendix E provides a listing of these factors. It is clear that technical and organizational factors encompass most of the empirical findings. The present study addressed all the factors proposed by the reviewers but focused on the area where organizational and technical factors blend, i.e., the content of the evaluation findings and

Figure 1

Mean Scores for Reports A-K



A=1.27, s.d.=1.Ø8
B=1.62, s.d.=1.93
C=2.58, s.d.=2.29
D=4.17, s.d.=2.74
E=4.23, s.d.=2.6Ø
F=4.66, s.d.=2.94
G=5.22, s.d.=3.48
H=6.39, s.d.=1.99
I=6.49, s.d.=3.67
J=7.Ø2, s.d.=3.52
K=7.23, s.d.=2.87

$\bar{X}$=4.64, s.d.=3.22

recommendations, which is beginning to be recognized as another important factor in the study of utilization.

The case studies

A case study approach was selected to address the second objective of this research because it was best suited to the level of specificity and quantity of detail that was necessary to address the purpose of this study, i.e., to compile a composite of characteristics of evaluation reports and the evaluation process associated with varying degrees of use of evaluation information. Concerning data collection and analysis for case studies, the methods included structured interviews with the primary audience and evaluators, surveys of secondary audiences, classification of content of evaluation information, and documentation of organizational actions and procedures common to all reports. The analysis produced a description of the relationship of the variables with utilization through the contrast of high-use and low-use reports.

Results of case studies

The results confirm that relevance to decision-making is a major factor influencing utilization in this context. Other variables clearly supported by this study included political and organizational circumstances, focus on manipulable variables, and evaluation user characteristics. Less consistent results were found for user involvement in study formulation, credibility of information, evaluator credibility in terms of program knowledge, and quality as important variables.

The findings also indicated that content of evaluation information is an important factor to consider in investigations of utilization. Recommendations in high-use studies were less variable in content compared to recommendation in low-use studies, and they tended to focus on program eligibility issues, service improvements, or improvements in management. Surprisingly, the findings suggest that policy-oriented recommendations and recommendations requiring interprogram or interagency action, even though harder to implement, had more influence on the decisions to implement the recommendations than the less challenging recommendations requiring only action by program managers.

Conclusions

The measurement process investigated in this study worked well as a measure of utilization of evaluation information, effectively differentiating between high- and low-use reports. Additional conclusions arising from the study address issues related to measurement of utilization. First, the results of the study provide evidence that judging utilization by looking at implementation of evaluation recommendations alone may be misleading. Evaluation recommendations may be implemented, but the decision to implement may have been made on the basis of entirely separate information. Studies which do not take influence into account could present an inflated picture of the extent to which evaluation information is used. Second, the study highlighted the fact that secondary audiences of evaluation studies should be included in an assessment of utilization. Finally, cognitive and persuasive uses of evaluation information were documented in this study, reinforcing the current assertion in the literature that the

definition of utilization should include conceptual and persuasive use in addition to instrumental use.

This study provides empirical evidence by which to judge the largely theoretical or anecdotal literature on utilization. The results of this investigation support some variables and do not support other variables found in the literature. No single listing of the variables in the five review publications fared significantly better than the others. However, evidence from this study provides validation for some of the commonalities among the listings.

Questions or comments concerning this paper should be directed to:

Dr. Nina B. Barrios
Evaluation Systems Design, Inc.
700 N. Calhoun St. Suite A-3
Tallahassee, Florida 32303

REFERENCES

Barrios, N.B. (1986). Utilization of evaluation information: A case study approach investigating factors related to evaluation utilization in a large state agency. Unpublished doctoral dissertation, Florida State University, Tallahassee.

Beyer, J.M. and Trice, H.M. (1982). The utilization process: A conceptual framework and synthesis of empirical findings. Administrative Science Quarterly, 27, 591-622.

Braskamp, L.A. and Brown, R.D. (1980). Utilization of evaluation information. San Francisco, AC: Jossey-Bass Inc., Publishers.

Guilford, J.P. (1954). Psychometric Methods. New York: McGraw-Hill Book Company.

Heilman, J.G. (1983). Beyond the technical and bureaucratic theories of utilization: Some thoughts on synthesizing reviews and the knowledge base of the evaluation profession. Evaluation Review, 7 (6), 707-728.

Leviton, L.C., and Hughes, E.F. (1981). Research on the utilization of evaluations: A review and synthesis. Evaluation Review, 7 (4), 525-548.

Patton, M.Q. et al. (1977). In search of impact: An analysis of the utilization of federal health evaluation research. In C.H. Weiss, Using social research in public policy making (pp. 141-164). Lexington, MA: Lexington, D.C. Health and Company.

Rich, R.F. (1981). Social science information and public policy making. San Francisco: Jossey-Bass Publishers.

Rothman, J. (1980). Using research in organizations: A guide to successful application. Beverly Hills, CA: Sage Publications.

Weiss, C.J. (1972). Evaluating action programs: Readings in social action and education. Boston: Allyn and Bacon, Inc.

Weiss, C.J. (1981). Measuring the use of evaluation. In J.A. Ciarlo (Ed.), Utilizing evaluation: Concepts and measurement techniques (pp. 17-31). Beverly Hills, Ca: Sage Research Progress Series in Evaluation, 6.

Weiss, C.H., and Bucuvalas, M.J. (1980). Social science research and decision making. New York: Columbia University Press.

Young, C.J. and Comptois, J. (1979). Increasing congressional utilization of evaluation. In F. Zweig (Ed.), Evaluation in legislation. Beverly Hills, CA: Sage Publications.

Appendix A

Equal Interval Scale Development:

Category Definitions and Response Form

SCALE DEVELOPMENT
EFFECTIVENESS OF EVALUATION RECOMMENDATIONS
CATEGORY DEFINITIONS

The attached scale is designed to develop an equal interval scale indicating effectiveness of evaluation recommendations. The terms listed here are used in the scale. The eleven point scale is anchored at each end by the terms "high effectiveness" and "low effectivenss", defined below. The scale items are pairs of the two categories, implementation and influence, also defined below.

Effectiveness

Low effectiveness--No utilization of information; No influence on decisions; Recommendation will not be implemented.

High effectiveness--Maximum utilization; Predominant influence on decisions; Recommendation will be implemented as stated.

Implementation Status

Implemented as stated--Recommendation was implemented with no modifications.

Implemented with modifications--Recommendation was implemented but in a modified form.

Planned implementation as stated--Recommendation will be implemented with no modifications.

Planned implementation with modifications--Recommendation will be implemented but in a modified form.

Not to be implemented--Decision has been made that the recommendation will not be implemented.

Decision not yet made--No decision regarding the implementation of the recommendation has been made.

Influence

Predominant influence--Decision probably would not have occurred if the evaluation results and recommendations had not been made; OR Prior decisions were recalled (cancelled) due to evaluation results and contadictory recommendations.

Major influence--Many factors were involved in the decision, but the evaluation results and recommendations weighed heavily in the final analysis; OR Prior decisions were postponed due to evaluation evidnece contradicting earlier consensus.

Moderate influence--The evaluation results and recommendations were about equally weighted with other factors in making the decision.

Minor influence--The decision probably would have occurred without the evaluation results and recommendations, but the information was recognized as supportive of the decision; OR A contradictory decision was made but the position of the evaluation was recognized as not lending support (considered and rejected.)

No influence--The evaluation results and recommendations were given no consideration in the process of decision-making.

SCALE DEVELOPMENT
EFFECTIVENESS OF EVALUATION RECOMMENDATIONS

Position:_____ Name (optional):_____

Directions: The situations listed down the left side describe the status of an evaluation recommendation and the extent of influence the evaluation had on the decision or action. (See attached page for a detailed description of the categories). The continuum across the top indicates a measure of effectiveness of evaluation information, from no effectiveness to complete effectiveness. For each situation, mark the point on the scale corresponding to your perception of its value.

	Effectiveness										
	Low 1	2	3	4	5	6	7	8	9	10	High 11
Implemented as stated / Predominant influence											
Implemented as stated / Major influence											
Implemented as stated / Moderate influence											
Implemented as stated / Minor influence											
Implemented as stated / No influence											
Implemented with modifications / Predominant influence											
Implemented with modifications / Major influence											
Implemented with modifications / Moderate influence											
Implemented with modifications / Minor influence											
Implemented with modifications / No influence											
Planned implementation as stated / Predominant influence											
Planned implementation as stated / Major influence											
Planned implementation as stated / Moderate influence											
Planned implementation as stated / Minor influence											
Planned implementation as stated / No influence											
Planned implementation w/modification / Predominant influence											
Planned implementation with modifications / Major influence											
Planned implementation w/ modifications / Moderate influence											
Planned implementation with modifications / Minor influence											
Planned implementation with modifications / No influence											
Not implemented / Predom. influence (Prior decision recalled)											
Not implemented / Major influence (Prior decision postponed)											
Not implemented / moderate influence											
Not implemented / minor influence											
Not implemented / no influence											
Decision not made / Predom. influence (Prior decision recalled)											
Decision not made / Major influence (Prior decision postponed)											

Appendix B

Scale Values and Measures of Dispersion (Q)

Results of Equal Interval Scaling Procedure
(Number of Respondents = 15)

Scale Items	Scale Values	Q Values	Frequency*
Implemented as stated/ Predominant influence	10.93	.57	1
Implemented as stated/ Major influence	9.99	.57	2
Implemented as stated/ Moderate influence	7.70	2.53	13
Implemented as stated/Minor influence	4.65	2.21	12
Implemented as stated/No influence	1.75	1.25	15
Implemented with modifications/Predominant influence	9.84	1.32	2
Implemented with modifications/Major influence	8.77	1.21	3
Implemented with modifications/Moderate influence	6.99	1.22	9
Implemented with modifications/Minor influence	4.09	2.38	6
Implemented with modifications/No influence	1.61	1.55	5
Planned implementation as stated/Predominant influence	9.73	1.44	2

*Number of times mentioned by primary audience in ratings of the 74 recommendations.

Planned implementation as stated/Major influence	9.Ø2	1.95	4
Planned implementation as stated/Moderate influence	6.4Ø	1.8Ø	6
Planned implementation as stated/Minor influence	4.15	2.79	4
Planned implementation as stated/No influence	1.65	1.61	1
Planned implementation with modification/Predominant influence	8.73	2.94	3
Planned implementation with modification/Major influence	7.73	3.21	4
Planned implementation with modification/Moderate influence	5.87	3.29	2
Planned implementation with modification/Minor influence	3.35	2.55	7
Planned implementation with modification/No influence	1.33	1.33	1
Not implemented/Predominant influence (Prior decision recalled)	6.89	4.68	1
Not implemented/Major influence (Prior decision postponed)	5.49	3.56	Ø
Not implemented/Moderate influence	4.27	5.76	Ø
Not implemented/Minor influence	2.25	3.Ø2	Ø

Not implemented/No influence	1.04	.54	12
Decision not made/Predominant influence (Prior decision recalled)	8.07	3.95	1
Decision not made/Major influence (Prior decision postponed)	6.21	3.19	0

Appendix C

Questionnaire for Primary Audience (Part A)

UTILIZATION OF EVALUATION INFORMATION

General Instructions:

The attached questionnaires are part of a study of utilization of evaluation information. The study will identify characteristics of the evaluation process and reports associated with high use studies and those associated with low use studies. The results of these questionnaires will help to make the initial rating of reports in terms of extent of use. The focus of the study is selected evaluation reports produced by the Office of Evaluation and Management Review during 1984 and 1985.

I am requesting that you complete the questionnaires in terms of the uses, if any, you or your office made of the report(s) specified on the forms. Part A addresses the status of the recommendations contained in the report and the extent of influence of the evaluation information on the decision. Part B addresses other uses of the evaluation information including conceptual and persuasive uses.

Please hold the completed questionnaires for personal pick-up by Nina Barrios. A completion date of January 6, 1986 is requested.

Although this study is being conducted privately (as a doctoral dissertation), it is being conducted with the permission and encouragement of , Inspector General and Director, Office of Evaluation and Management Review. General findings will be shared with the Inspector General's office while specific results, such as you response to these questionnaires, will be confidential.

Your assistance and cooperation is greatly appreciated.

UTILIZATION OF EVALUATION RESULTS

Part A - Code Explanations

Implementation Status

Implemented as stated--Recommendation was implemented with no modifications.

Implemented with modifications--Recommendation was implemented but in a modified form.

Planned implementation as stated--Recommendation will be implemented with no modifications.

Planned implementation with modifications--Recommendation will be implemented but in a modified form.

Not to be implemented--Decision has been made that the recommendation will not be implemented.

Decision not yet made--No decision regarding the implementation of the recommendation has been made.

Influence

Predominant influence--Decision probably would not have occurred if the evaluation results and recommendations had not been made; OR Prior decisions were recalled (cancelled) due to evaluation results and contadictory recommendations.

Major influence--Many factors were involved in the decision, but the evaluation results and recommendations weighed heavily in the final analysis; OR Prior decisions were postponed due to evaluation evidnece contradicting earlier consensus.

Moderate influence--The evaluation results and recommendations were about equally weighted with other factors in making the decision.

Minor influence--The decision probably would have occurred without the evaluation results and recommendations, but the information was recognized as supportive of the decision; OR A contradictory decision was made but the position of the evaluation was recognized as not lending support (considered and rejected.)

No influence--The evaluation results and recommendations were given no consideration in the process of decision-making.

STATUS AND INFLUENCE OF EVALUATION RECOMMENDATIONS

REPORT:_____ RESPONDENT:_____

DIRECTIONS: In the box on left, enter code reflecting current status of recommendation. In box on right, enter code reflecting the influence of the evaluation results on the decision.

STATUS CODES	INFLUENCE CODES
(1) Implemented as stated	(1) Predominant influence
(2) Implemented with modifications	(2) Major influence
(3) Planned implementation as stated	(3) Moderate influence
(4) Planned implementation w/ modifications	(4) Minor influence
(5) Not to be implemented	(5) No influence
(6) Decision not yet made	

☐	1. In view of the small fraction (3%) of cases at risk in which intervention has occurred, the number and proportion of cases in which intervention is made should be increased. This goal may be attained more easily if additional modes of intervention (notification procedures) are used, including client education.	☐
☐	2. To determine the usefulness of information generated by the DUR Project to peer review and case management, PDDM should establish formal links between the DUR Project and the other quality control systems, including means to provide information from DUR to peer review groups.	☐
☐	3. In view of the equivocal character of present evidence of the project's effects, PDDM, in cooperation with HID, Inc., should conduct further pre/post analyses upon individuals subject to direct intervention, such as comparing risks among patients of physicians who have received the letters with changes in risk among others for whom profiles were generated, but letters were sent, and those not at risk.	☐
☐	4. If, as appears, the majority of cases at risk involve a limited number of risky combinations of drugs and diagnoses, then the project should intensify and broaden efforts to inform all Medicaid physicians concerning these risks and to encourage patient education efforts.	☐

REPORT:_____ PAGE___ OF ___

STATUS CODES

(1) Implemented as stated
(2) Implemented with modifications
(3) Planned implementation as stated
(4) Planned implementation w/ modifications
(5) Not to be implemented
(6) Decision not yet made

INFLUENCE CODES

(1) Predominant influence
(2) Major influence
(3) Moderate influence
(4) Minor influence
(5) No influence

☐ 5. Improvements should be made to the reference material provided to committee members. Specifically, the United States Pharmacopeial Dispensing Information, a reference volume which is highly rated by those committee members who use it, should be made available to all members. In addition, FPA should continue to provide information drawn from the current pharmacological literature. ☐

☐ 6. As an alternative to no further action on cases for which the profile contains insufficient information, or reflects a risk considered too insignificant to necessitate an inquiry, the committees should be given the option of making a referral to a Medicaid case manager who could assure that either more information is obtained for committee review, or that the client is counselled on the appropriate use of prescribed medications, etc. ☐

Appendix D

Questionnaire for Primary Audience (Part B)

UTILIZATION OF EVALUATION INFORMATION
Part B

Report:_____ Respondent:_____

Instructions: Please complete the following questions in terms of the report specified above.

1. Were any decisions made or actions taken as a result of evaluation information prior to the completion of the report (and therefore not included in recommendations)? Briefly describe.

2. Were any decisions made or actions taken as a result of evaluation information though not specifically recommended in the report? Briefly describe.

3. To what extent did the evaluation information contribute to a general view or understanding of the program?

Not at all Little Some Much

4. To what extent was the evaluation information used as a reference in further evaluation?

Not at all Little Some Much

5. To what extent did the evaluation information stimulate new activities?

Not at all Little Some Much

6. To what extent did the evaluation information convince you of the need for new action, practice or policy?

Not at all Little Some Much

7. To what extent did the evaluation information convince you to maintain or intensify existing actions, practices or policies?

Not at all Little Some Much

8. Have you used the evaluation information to convince others to take action or make a specific decision? If so, briefly describe the situation(s).

9. Has the evaluation information been used by you in any other way, not addressed in Part A or B of this questionnaire? If so, briefly describe.

Appendix E

Organization of Variables for Interviews and Analysis

Technical

Formulation:

Relevance of information, decision driven
Focus on manipulable variables
Focus on process as well as outcome
User participation

Production:

Quality (including use of appropriate methods)
Timeliness
User participation
Style and credibility of evaluator
Credibility of information
Report characteristics--readability, jargon, clarity, style, explicit recommendations

Communication of findings:

Periodic reporting of findings
How disseminated--verbal, written
To whom disseminated--all relevant parties
Evaluator role--educator, advocate

Organizational

User characteristics--resistance/commitment to study, advocate of information, preconceptions toward evaluation
Organizational characteristics--linking mechanisms, relationship between units, values of organization, organizational capability, communication patterns
Prevailing circumstances in terms of specific study, including political circumstances
Evidence of "different cultures" phenomenon

UTILIZATION OF EVALUATION INFORMATION:

A case study approach investigating factors related to evaluation utilization in a large state agency

Authors: Nina B. Barrios, Ph.D., Evaluation Consultant, Evaluation Systems Design, Inc. and Garrett R. Foster, Ph.D., Professor, Florida State University

Address/Phone: Evaluation Systems Design, Inc. 700 N. Calhoun St. Suite A-3, Tallahassee, FL 32301. Phone: 904-681-6553.

ABSTRACT

This investigation measured evaluation utilization in a large state agency and used a case study approach to investigate conditions that related to utilization of evaluation reports. One significant contribution of this study was development of a strategy to measure influence of evaluation information on decisions and implementation status of recommendations. The measurement strategy produced scale values for reports, enabling a comparative rating of evaluation reports in terms of utilization. The ratings identified high-use and low-use reports.

A method for measuring utilization was used which: (1) incorporated the three-part definition of utilization presently accepted in the literature (instrumental, conceptual and persuasive), (2) looked at instances of recommended and unprompted actions, (3) accounted for competing influence on decision-makers and (4) allowed for quantification for judgment and comparison purposes. Additionally, the measurement strategy incorporated information from primary and secondary audiences of the evaluation reports.

To measure instrumental utilization, a scaling procedure was developed that addressed both the status of the recommendations within a report and the extent of influence of the evaluation of the status. The scaling procedure, the equal-appearing interval method, used comparative judgments from 15 experienced evaluators to obtain scale values for statements reflecting the relative importance of varying combinations of status and influence in the overall assessment of the effectiveness of evaluation information. The scale was tested on 11 evaluation reports produced by internal evaluators in a large state human services agency. Each recommendation in the 11 reports was assigned a scale value according to the results of the equal-appearing interval scaling procedure. To score reports, the scores for each recommendation were totaled and the mean and standard deviation computed for each report.

In addition to the ratings reflecting instrumental uses of evaluation information, members of the primary audience were also asked to respond to a questionnaire reflecting forms of utilization other than instrumental.

Application of the scaling process to the 11 reports allowed selection of four reports for indepth case studies which investigated the conditions that relate to utilization of evaluation reports in a large state agency. The four reports selected for study illustrated situations where the extent of utilization was high and three where utilization was low. The contrast of high-use and low-use reports provided a basis for assessing the potency of various utilization predictors (factors) reported in the literature.

Overall, the scale appeared to function as intended, i.e., to provide meaningful scale values that discriminate between the varying levels of utilization, defined as the combination of status and influence. Results of the case studies confirm that relevance to decision-making is a major factor influencing utilization in this context. Other variables clearly supported by this study included political and organizational circumstances, focus on manipulable variables, and evaluation user characteristics. Less consistent results were found for user involvement in study formulation, credibility of information, evaluator credibility in terms of program knowledge, and quality as important variables. Findings also indicated that content is an important factor to consider in investigations of utilization. Several notable findings are presented with regard to relationship of content of recommendations to utilization.